CONTROLLED SERVER LOADING
USING L4 DISPATCHING

Abstract of the Disclosure

    Standalone and cluster-based servers, including Web
servers, control the amount of data processed concurrently
by such servers to thereby control server operating
performance. A dispatcher is preferably interposed between
clients and one or more back-end servers, and preferably
monitors the performance of each back-end server (either
directly or otherwise). For each back-end server, the
dispatcher preferably also controls, in response to the
monitored performance, either or both the number of
concurrently processed data requests and the number of
concurrently supported connections to thereby control the
back-end servers' performance. In one embodiment, the
dispatcher uses a packet capture library for capturing
packets at OSI layer 2 and implements a simplified TCP/IP
protocol in user-space (vs. kernel space) to reduce data
copying. Commercially off-the-shelf (COTS) hardware and
operating system software are preferably employed to take
advantage of their price-to-performance ratio.